DOCUMENT RESUME

ED 069 677 TM 002 128

AUTHOR Fiske, Donald W.; Kuncel, Ruth Boutin
TITLE How Does It Feel to Take A Personality Test?
SPONS AGENCY National Science Foundation, Washington, D.C.
NOTE 11p.; Solicited by the American Psychological Association

EDRS PRICE MF-$0.65 HC-$3.29
DESCRIPTORS Affective Tests; Attitude Tests; Catalogs; Evaluation Techniques; Measurement Techniques; *Personality Tests; Psychological Evaluation; Research Methodology; *Response Style (Tests); Student Attitudes; Technical Reports; *Testing; *Test Reliability; *Tests
IDENTIFIERS Marlowe Crowne Social Approval; Thurstone Dominance; Thurstone Reflective; Welsh Anxiety; Welsh Repression

ABSTRACT

After taking a personality test, subjects reported their reactions to being tested. Reactions were diverse, even in the same subject. Free responses to 10 questions were coded into 16 categories within five broad groups. Desire for information about the test and about self, and criticism of testing were very prevalent; criticisms of self and apprehensions were less frequent. Three studies yielded rather similar results. Inventories, however, elicit somewhat different patterns of reactions from other tests. These diverse reactions may underlie response sets and, consequently, impair the quality of personality measurements. Appropriate increases in the structuration of our tests seem called for. (Author)

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY

ED 069677

Abstract

How Does It Feel to Take a Personality Test?

Donald W. Fiske and Ruth Boutin Kuncel

University of Chicago

After taking a personality test, Ss reported their reactions to being tested. Reactions were diverse, even in the same S. Free responses to 10 questions were coded into 16 categories within five broad groups. Desire for information about test and about self, and criticism of testing were very prevalent; criticisms of self and apprehensions were less frequent. Three studies yielded rather similar results. Inventories, however, elicit somewhat different patterns of reactions from other tests.

These diverse reactions may underlie response sets and, consequently, impair the quality of personality measurements. Appropriate increases in the structuration of our tests seem called for.

FILMED FROM BEST AVAILABLE COPY

ED 069677

HOW DOES IT FEEL TO TAKE A PERSONALITY TEST?

Donald W. Fiske and Ruth Boutin Kuncel

University of Chicago

Of the two components in the person-test interactions used for personality measurement, the test has been investigated much more thoroughly than the person. The more limited research on the person (cf. Fiske, 1971, Ch. 10) has centered on effects of response sets, on reactions to individual items (Kuncel, in press) and on effects of general reactions (e.g., test anxiety) on test performance (cf. Wine, 1971). Following up an earlier survey (Fiske, 1967), this paper examines reactions to personality tests and to taking them, as reported by Ss after taking a test.

Method

Subjects were first given a personality test in standard form under anonymous conditions. The booklet continued with a statement that we were interested in how the subject feels about taking such tests. Then, three open-ended questions asked: "What did you think or feel when you found out...that you were being asked to take a test?" "What did you think or feel while you were taking it?" and "Did you have any other thoughts or feelings" while taking such tests on prior occasions? More specific questions followed: "When you take a personality test, what do you

want to know about it?" "What would you like to know about your answers to it?" "What do you dislike most about taking a personality test?" "What might make you feel uneasy" about taking a personality test? "When you have trouble deciding how to answer a question on a personality test..., what seems to cause the difficulty?" "What other comments can you give us on personality tests and on taking them?"

In the first study, the personality test consisted of 15 yes-no or true-false items selected unsystematically from several standard instruments. In the next study, the 40 items included ten from each of four scales: Thurstone Dominance, Thurstone Reflective, Welsh Repression, and Welsh Anxiety. In the last study, each S received one of three instruments: Marlowe-Crowne Social Approval, Thurstone Embedded Figures (21 items given without time limit), or Similes Preference Inventory (Pearson & Maddi, 1966). The last asks Ss to select the ending of the simile which he prefers, choosing among five ranging from the standard cliche to highly novel or absurd endings. This paper will report the last study and will indicate how its findings agreed or differed with those of the earlier ones. Each of the three studies employed men and women from a different university. College students were used because they are relatively verbal and introspective. Participation was semi-voluntary.

Sixteen reaction categories, induced from the free responses, were later grouped into five classes on the basis of similarity;

Desire for information about test - purpose of test or testing, who is the examiner and is he honest and competent, what is the whole test like, how were the data analyzed, what are the results; desire for information about self - desire to learn about self, where do I stand; criticisms of testing - tests are insufficiently valid, items or tests are uninteresting or repetitive or inconsistent or blatant, items are ambiguous; criticisms of self - lack required information about self, desire to be honest; apprehensions - fear of learning about self, fear of others learning about self, fear of being criticized, wish to have others know S as an individual.

Each response was rated for the clear presence, the possible presence, or the absence of the reaction categories. Only the coding for clear presence is reported here. Raters used a manual. After reasonable inter-rater agreement was achieved, one rater rated all protocols.

Insert Table 1 about here

Results

Table 1 gives the proportions of Ss taking each test whose responses were coded into each of the five broad categories. The occurrences for the unstructured initial three questions are given separately. Note that, in response to the unstructured

initial three questions, substantial proportions of Ss want more information about the test and are critical of testing. In their total protocols, high frequencies of all five categories are found. The patterns for Marlowe-Crowne and Similes are fairly similar in desire for information and in criticism of testing, the deviant pattern for Embedded Figures probably resulting from Ss' puzzlement over a test of competence being treated as a personality test (the concept of field dependence not yet being in the public domain). The patterns of Embedded Figures and Similes tend to differ from Marlowe-Crowne on criticism of self, probably because of the more threatening material contained in a test of social approval.

The sixteen categories yielded intercorrelations best interpreted as chance variations around zero. Few correlations with test scores, order of completion, and number of "true" responses on the Marlowe-Crowne were statistically significant.

In spite of the differences in populations sampled and in tests used, the findings of this third study agree closely with the prior two. The prevalences of the sixteen categories were similar with but three exceptions: this group manifested more

desire to learn about themselves and less feeling that tests were uninteresting, etc. Also, many more of the subjects taking personality inventories found them ambiguous (67%) than did subjects taking either the Embedded Figures or Similes test (15%).

Interpretations and Implications

Objectively, the test items, the instructions, and the conditions may be constant for all Ss, as comparative measurement requires (cf. Fiske and Butler, 1963). Subjectively, however, tests elicit a diversity of reactions within and between subjects, even under anonymous conditions. Our median S demonstrated 4 or 5 of the 16 categories. The diversity probably reflects in part the S's lack of a fixed style for handling such a test situation. Subjects are generally unfamiliar with tests; they have received little feedback from other testing situations; and consequently they have probably generalized from aptitude and achievement tests to any tests of personality. As a result of the lack of fixed style, we would guess that protocols did not report all of the reactions that the Ss experience. On the other hand, it must be recognized that, in our instrument, the later questions about their general experiences with personality tests were highly permissive if not leading. Our goal was to discover what qualitative reactions occurred, not to estimate their frequency or strength. It should also be noted that the questions Ss answered about their reactions were themselves a "test." We cannot say

to what extent their uncertainty about who would read their answers may have affected their responses.

This diversity of reactions observed in and between subjects was also observed between tests. In these studies, questionnaires aimed at self-perception (such as the Marlowe-Crowne) elicited twice as many reactions in the class labelled criticisms of self, twice as many spontaneous expressions of apprehensions, and slightly more criticisms of tests than did those of preference and ability (Similes and Embedded Figures). Perhaps, by directing attention inward, self-descriptive instruments generate more "self-evaluative, self-deprecatory thinking" and intensive anxiety about being tested (cf. Wine, 1971).

Although we have identified a series of reactions and shown that they vary within and between subjects as well as across tests, we can say little about their interrelationships or origins. The absence of important correlations in this study, as in the earlier two, probably results from our concentration upon the presence of varied reactions rather than differences in strength.

The very presence of these various reactions does, however, suggest an origin for many of the recognized response sets. For instance, the basis for the response set called social desirability is evident in the four fears we have subsumed under apprehensions and in the infrequently expressed concern with being honest. Furthermore, the tendency toward cautious, guarded

responding probably stem from uncertainties arising from a lack of information about the test, apprehensions about test outcome, etc.

In his original paper on response sets, Cronbach (1946) argued that response sets appear when S has difficulty in answering an item. Such difficulty is associated with items close to the subject's scale position (Runcel, in press). We would predict that, at these points, the S's varying reactions to being tested would have most influence.

The diversity of reactions to tests observed in these studies and in an earlier one (Fiske, 1967), may contribute to the substantial proportion of unsystematic variance in the matrix of test responses (cf. Fiske, 1971, Ch. 8). Perhaps some of these reactions and, thus perhaps the basis for some response sets, can be eliminated by providing the subject with increased information about the test, more feedback about the results, etc. To the extent that these reactions cannot be eliminated and that they are found to influence adversely test results, there remains the unsolved problem of designing operations for measuring personality which minimize their effects. Extending the recommendation in Cronbach (1946) and the similar recommendation in Fiske (1968), we must seek some method of appropriately increasing the structuration of our tests.

References

Cronbach, L. J. Response sets and test validity. Educational and Psychological Measurement, 1946, 6, 475-494.

Eisenberg, P., & Wesman, A. G. Consistency in response and logical interpretation of psychoneurotic inventory items. Journal of Educational Psychology, 1941, 32, 321-338.

Fiske, D. W. Some hypotheses concerning test adequacy. Educational and Psychological Measurement, 1966, 26, 69-88.

Fiske, D. W. The subject reacts to tests. American Psychologist, 1967, 22, 287-296.

Fiske, D. W. Measuring the concepts of personality. Chicago: Aldine, 1971.

Fiske, D. W., & Butler, J. M. The experimental conditions for measuring individual differences. Educational and Psychological Measurement, 1963, 23, 249-266.

Kuncel, R. B. Response processes and relative location of subject and item. Educational and Psychological Measurement, in press.

Pearson, P. H., & Maddi, S. R. The Similes Preference Inventory: Development of a structured measure of the tendency toward variety. Journal of Consulting Psychology, 1966, 30, 301-308.

Wine, J. Test anxiety and direction of attention. Psychological Bulletin, 1971, 76, 92-104.

Footnote

[1]This research was conducted under Grant No. GS-3127 from the National Science Foundation. We are also indebted to D. Susan Butt for her assistance in gathering the data, to Dale Larson for making the ratings, and to Michael Minor for carrying out analyses.

Table 1

Observed Reactions to Three Tests

(Percents of Ss Displaying Each Class of Reaction)

	Three initial unstructured questions			Total free-response protocol		
	M-C	EF	Sim.	M-C	EF	Sim.
Desire for information about test	24	50	24	70	100	76
Desire for information about self	27	04	24	85	54	84
Criticisms of testing	36	17	36	76	46	60
Criticisms of self	30	13	08	52	21	20
Apprehensions	24	08	12	76	83	56
(N)	(33)	(24)	(25)	(33)	(24)	(25)